Why Only Externalists Can Be Steadfast

de Ridder, G.J.

published in
Erkenntnis
2013

DOI (link to publisher)
10.1007/s10670-013-9452-5

document version
Early version, also known as pre-print

Why Only Externalists Can Be Steadfast

Jeroen de Ridder

Faculty of Philosophy, VU University Amsterdam

De Boelelaan 1105

1081 HV Amsterdam

The Netherlands

+31 20 5984161

g.j.de.ridder@vu.nl

Abstract

What is the rational response to disagreement with an epistemic peer? Some say the steadfast response of holding on to your own belief can be rational; others argue that some degree of conciliation is always rationally required. I argue that only an epistemological externalist about rationality — someone who holds that the rationality of a belief is partly constituted by factors outside a subject's cognitive perspective — can defend the steadfast view. Or at least that this is so in the kinds of idealized cases of peer disagreement that take center stage in the current debate about disagreement. The argument has three steps. First, I show how rationality internalism motivates conciliationism: In view of the mutually recognized internal epistemic symmetry between peers, it would be arbitrary for either peer to hold on to her own belief. Second, I strengthen this line of thought by considering various proposed 'symmetry breakers' that appear to introduce a relevant asymmetry between peers, which could be used to defend the rationality of a steadfast response. I argue that none of these alleged symmetry breakers can help internalists. Third, I show how externalism does have the resources to

defend steadfastness and expose how extant defenses of steadfastness implicitly rely on externalist intuitions.

1. Introduction

The problem of disagreement is familiar by now. What is the rational response when you find out that an epistemic peer disagrees with you about whether p? Proponents of conciliationism argue that you should move your credence in the direction of your peer or give your belief up. Proponents of steadfastness argue that you can be rational in sticking with your belief.

A striking feature of the debate about disagreement is the virtual absence of attempts to flesh out the relevant notion of rationality in more detail.[1] The typical methodology is to present cases and to rely on intuitive judgments about what is rational in them. This is unfortunate, because how you understand rationality matters a great deal for what you can say about the rational response to peer disagreement. I will argue that only rationality externalists can defend steadfast views, at least in idealized cases of peer disagreement.

In section 2, I clarify what I mean by idealized cases of peer disagreement. Section 3 provides characterizations of rationality internalism and externalism. In sections 4 and 5, I argue that rationality internalists lack the resources to defend a steadfast view. In section 6, I show that externalism does have the resources to defend a steadfast view.

2. Idealized Disagreement

The epistemological discussion tends to focus on what we can call *idealized disagreement* (cf. Lackey 2010: 303): known disagreement with a genuine epistemic peer, who is also recognized as such. It is doubtful whether what we say about such idealized cases provides

guidance for *real-life disagreement*. First, genuine epistemic peers may be rare or even nonexistent. Following Kelly (2005, pp. 174–5), peers are typically understood as each others' equals with respect to (i) familiarity with the relevant evidence and arguments bearing on the issue at hand and (ii) general epistemic virtues such as intelligence, thoughtfulness, and freedom from bias.[2] Various aspects of peerhood appear problematic in real life (King 2011; Frances 2010a, pp. 424–5). Perhaps each of us has a unique combination of epistemic virtues and vices that influences how we form our beliefs. Perhaps our evidence contains hard-to-share seemings, insights, or feelings (Van Inwagen 1996, 2010; Sosa 2010). Or perhaps it is partly constituted by first-personal assessments and weightings of our sources of evidence. If so, then no one is ever anyone's peer in a strict sense. Second, even if genuine peerhood is possible, perhaps we couldn't recognize it. It may be that sharing all your evidence, knowing that both of you have done so, and verifying that you have understood each other correctly is unattainable, practically or even in principle.

When I claim that only externalists can be steadfast, I mean that externalism has the resources to defend a steadfast view *in idealized disagreements*, whereas internalism does not. It does not follow from this that internalists cannot defend a steadfast response in real-life disagreements. Since real-life disagreements rarely if ever meet the strict conditions of idealized disagreements, it will typically be possible to appeal to circumstances or factors of the disagreement that differ between you and your peer. Such asymmetries may then enable a steadfast response.

My conclusion thus has limited implications for what internalists can say about real-life disagreements. At most it entails that, the more a real-life case resembles the idealized scenario, the harder it will be for internalists to defend a steadfast response. Nonetheless, it is an important theoretical result that there is a connection between the internalism-externalism divide in epistemology and views about the rational response to disagreement.

3. Internalism and Externalism

In the disagreement literature 'rational' is used interchangeably with 'justified' (White 2005, p. 445; Feldman 2006, pp. 220f; Kelly 2010; Ballantyne & Coffman 2011).[3] I'll follow suit here.

Although internalism is characterized in different ways by different authors, there is widespread agreement that its core is the idea that all the factors that determine the rationality (justification) of a belief must be *internal to the subject's cognitive perspective*, in the sense that they must accessible to her upon reflection. BonJour (2010a, p. 364) states:

> The most generally accepted account of this distinction is that a theory of justification is *internalist* if and only if it requires that all of the factors needed for a belief to be epistemically justified for a given person be *cognitively accessible* to that person, internal to his cognitive *perspective*.[4]

Rationality-conferring factors can include things such as the grounds for the belief and their adequacy, evidence for its truth and the strength thereof, the belief's being appropriately based on adequate grounds, or believed in view of the relevant evidence.

Externalism is the denial of internalism. It is the thesis that it is *not* the case that all the factors that determine the rationality of a belief must be internal to the subject's cognitive perspective and accessible to her upon reflection. The negative characterization allows for different versions of externalism, depending on how many and which rationality-conferring factors are external to the subject's cognitive perspective. Examples of external rationality-conferring factors include: adequacy of the grounds for your belief (Alston 1988), reliability

of your belief-forming processes (Goldman 1979), proper functioning of your cognitive system (Bergmann 2006), virtuousness of your belief-formation (Sosa 2007; Greco 2010), or social relations in which you stand (Meeker 2004).

4. From Internalism to Conciliationism

Consider an idealized case of disagreement: Alice and Beth are epistemic peers and know this to be so. They disagree about whether p: Alice believes p and Beth believes not-p (or something that obviously entails it). What's more, (i) both Alice and Beth are intelligent and epistemically responsible believers, and (ii) p is a proposition in their general area of competence; p is not something about which they are uninformed or ill-equipped to think about.[5]

In such a case, there's a strong initial intuition to say that because their situations are entirely symmetrical in terms of epistemically relevant factors both Alice and Beth should adjust their respective beliefs in each other's direction, and presumably withhold judgment about whether p. Considerations about epistemic symmetry underlie most of the defenses of conciliationism. Feldman, for instance, says:

> In situations of full disclosure, where there are not evident asymmetries, the parties to the disagreement would be reasonable in suspending judgement about the matter at hand. (Feldman 2006, p. 235)

According to Christensen, the main motivation for conciliationism is:

> that the peer's disagreement gives one evidence that one has made a mistake in interpreting the original evidence, and that such evidence should diminish one's confidence in P. (Christensen 2009, p. 2)

Note that this is true for both peers: Both acquire evidence suggesting that they made a mistake and hence should diminish confidence in their beliefs. The thought is that, since there's nothing available to the disagreeing peers to suggest an asymmetry between them with regard to how well-positioned they are to discern the truth or falsity of p, each should become doubtful about her belief and adjust her credence in the direction of the other or give up the belief.[6]

The first step in my argument is to note that these symmetry considerations are most compelling when understood along internalist lines: Everything that is *cognitively accessible* to Alice and Beth, both immediately and upon reflection, is wholly symmetrical. For all Alice and Beth can tell, the following things are true:

- We are both intelligent people thinking about an issue about which we are competent.
- Beth/Alice is my epistemic peer; she's equally intelligent, adept at reasoning, etc. She is equally well-informed as I am and familiar with the evidence and arguments bearing on p.
- We have disclosed and shared all our evidence.
- We have verified whether we understand each other correctly on every relevant point.

Given that both Alice and Beth hold these things to be true, or even know them to be true, everything that is relevant to rationally believing p or not-p really does seem to be symmetrical from Alice and Beth's cognitive perspectives. Hence, neither one of them will

have cognitive access to a 'symmetry breaker'[7] which would allow them to discount the other's opinion. For both Alice and Beth, sticking to their own belief seems arbitrary and thus irrational, just like it would be arbitrary to trust one of two thermometers with different readings, when you have no reason to assume that one is more reliable than the other.[8] At first pass, then, it seems that rationality internalism indeed leads to conciliationism, by means of the powerful intuition that mutually recognized internal epistemic symmetry should motivate both peers to move their beliefs in each other's direction.

5. No Internal Symmetry Breakers

Perhaps, however, the abstract characterization above overlooks relevant factors that might break the symmetry between Alice and Beth. One reason to take this possibility seriously is that although proponents of conciliationism appeal to symmetry in defending their view, they often deny that their view leads to widespread agnosticism. They think the symmetry can be broken in many cases. Therefore, the second step of my argument is to investigate whether closer inspection reveals symmetry breakers that allow internalists to defend a steadfast response in idealized disagreements after all.

Let me set aside a few red herrings first. Peers may become aware of special circumstances that negatively affect the other party's ability to assess the evidence, reason, or understand the relevant considerations on this particular occasion. All sorts of things could be relevant: physical or psychological well-being, disturbing factors in the environment that someone is sensitive to, a blind-spot with regard to the specific proposition at hand, etc.[9] This may happen, but when it does, we're no longer dealing with a case of idealized disagreement. Instead, in such a case Alice and Beth cease to be peers, even though they generally are.

Another possibility is that the social-epistemic environment affects Alice and Beth's disagreement. If Alice finds out that, say, 10 further peers independently agree with her about p while no peer thinks that not-p, or that a substantial majority of experts believe that p, she acquires a strong reason to hold on to her belief that p. Beth should become much less confident that not-p and perhaps should start believing that p instead.[10] (Recall that in an idealized disagreement Alice and Beth have already shared all their evidence, so it cannot be the case that Beth has access to an especially powerful piece of evidence that would sway both Alice and the other peers or experts.) This scenario, too, is not a case of idealized disagreement. It may have been one before the evidence about the social consensus came in, but once Alice and Beth become aware of the consensus it no longer is.

What about other possible symmetry breakers? Kelly (2005, pp. 178–80) suggests that the mere fact that Alice and Beth disagree about p could be a relevant symmetry breaker. Alice can think that she is right and that, therefore, Beth must have made a mistake on this particular occasion (and Beth can think likewise about Alice). This is consistent with their also continuing to regard each other as genuine peers and as having shared all their evidence. However, this response would be arbitrary from both their perspectives (e.g., Christensen 2007; Elga 2007).[11] Their disagreement gives Alice and Beth higher-order evidence that they may have made a mistake. In view of the mutually recognized symmetry, each should realize that she is just as likely as her peer to have made a mistake. The mere fact that it is *she herself* who believes p doesn't give Alice a good reason to think she, rather than Beth, is right (and likewise for Beth).

It is tempting to draw a general conclusion from this: An acceptable symmetry breaker must always be independent of the reasoning or other belief-forming process by which the peers initially formed their beliefs.[12] However, both Kelly (2010) and Lackey (2010) offer apparent counterexamples to this independence constraint. For instance, in cases where our

peer reports a belief that strikes us as utterly absurd or beyond the pale, we do want to allow this assessment of ours to function as a symmetry breaker, even though it does not seem to be independent of how we originally formed our belief.[13] Whether or not some general independence constraint ultimately stands up to scrutiny need not concern us here, however, because the point at issue is whether Alice could break the symmetry by simply relying on her own belief-forming process to discount Beth's contrary belief. Surely this sort of direct dependence on Alice's original belief-formation is unacceptably question-begging when Alice recognizes that Beth is her peer and is hence equally well-positioned to form a correct belief as to whether p.

Perhaps internalists could point to subjective or first-personal aspects of evidence to break the symmetry. Some evidence could be irreducibly first-personal (e.g., *your* feeling that *you* are in pain) or evaluations of publicly available evidence could have an irreducibly first-personal quality[14]. Since first-personal phenomena are accessible from within the subject's cognitive perspective, they can count towards the rationality of beliefs on an internalist understanding of rationality. Nonetheless, it would be hard or impossible to share them because of their first-personal character. Although subjects can attempt to share their first-personal perspectives with each other, they cannot make one another *have* them.

Given that we're considering idealized cases of peer disagreement, it cannot be the case that either Alice or Beth has significantly different irreducibly first-personal evidence, for that would violate the definition of peerhood. (There couldn't be genuine *peer* disagreement about whether, say, Alice has a headache, since Alice would possess crucial first-personal evidence that Beth lacks and hence there would be a significant evidential asymmetry between them.) Having different first-personal evaluations of the evidence would seem to be compatible with peerhood, though — at least, that is what almost most participants in the debate assume, as we shall see presently.[15]

I will now argue that appeals to a first-personal perspective to break the symmetry cannot be successful. I'll first give a general characterization of the problem and then we'll look at how some examples from the literature suffer from it. The problem is this: Given the mutually recognized symmetry of Alice and Beth's situations, for either one of them to trust their own first-personal perspective over that of the other seems arbitrary, for the same reasons as given in the previous section. There's nothing *epistemically* special about a first-personal perspective that warrants trusting it more than a recognized peer's perspective. So in the absence of reasons for thinking that their own perspective is epistemically superior, both Alice and Beth acquire higher-order evidence that their perspective may be misleading and should therefore become hesitant to continue relying on it. (And note that if there were such reasons, Alice and Beth — being peers — would have already shared them. Given that we're assuming that there still is disagreement, however, they must have evaluated these reasons differently.) Alice will realize that Beth, too, has a first-personal perspective which makes *her* think *she* is right. Their internal situation is again symmetrical, so sticking to their own belief would be arbitrary and hence irrational (and likewise for Beth).

One could try to resist this objection by arguing that there is a good reason to trust one's own first-personal perspective more than that of another. A suggestion from Wedgwood (2010, pp. 239–243) about trusting one's own moral intuitions can be adapted in this direction. Wedgwood claims that it can be rational for me to trust *my* moral intuitions rather than your conflicting intuitions, because only my intuitions can play the role of *directly* guiding my moral belief-formation. *Your* intuitions cannot. At best, they guide my belief-formation indirectly, through my beliefs about them. By analogy, one could propose that because only *my* first-personal perspective on the evidence can guide my belief-formation directly, it can be rational for me to trust it over yours. I am not convinced by this. Although the observation that someone's own intuitions or other first-personal experiences play a

special role by directly guiding one's belief-formation is in order, what remains unclear is why this fact should make it *epistemically* rational to trust my own first-personal perspective more than yours (given that we are peers). It may be practically rational or even inevitable to trust my own first-personal perspective, but that doesn't show that it is more likely to be right. To repeat the earlier point, disagreement with a peer gives you higher-order evidence that your first-personal perspective may be mistaken and that it could thus be wrong for you to keep trusting it, no matter how directly it guides your belief-formation.

Christensen (2007) and Elga (2007) also appeal to first-personal symmetry breakers. Christensen considers a case in which a peer reports a seemingly absurd belief about the size of your share of a restaurant check. The reason why it is not rational for you to move your belief towards your peer's absurd belief is that in this case the *best explanation* of your disagreement is that your peer has made the mistake, and not you. This is so, because the absurdity of her belief gives you a reason to think she, unlike you, has not employed commonsense checking, which is an extremely reliable procedure to check the correctness of simple mathematical beliefs (Christensen 2007, p. 201). I don't think this works. First of all, one may doubt whether this is even a possible example of peer disagreement. For it is hard to see how someone who calculates an absurdly high share of a bill and doesn't immediately check and correct her belief could be a peer of any reasonably educated and healthy adult. But, second, even if we grant that the case is possible, it's not clear that Christensen has identified a feasible symmetry breaker. The mere fact that someone is your peer already gives you evidence that she will have brought commonsense checking to bear on her calculation, since this is what normally functioning educated adults do. But if that isn't enough, we can amend the case so that your peer assures you in all sincerity you that she has used commonsense checking (cf. Bogardus 2009, p. 328–9).[16] Once you learn this, the symmetry

breaker is undermined and you no longer have a reason to think that the best explanation of your disagreement is that your peer has made a mistake.

Elga (2007, pp. 488–91) points to the "circumstances of the disagreement" as a source of potential symmetry breakers. I already identified several of these as red herrings near the beginning of this section, but in discussing a similar restaurant case as Christensen, Elga mentions one specific circumstance that may still be relevant here. It is your *extreme confidence* in your own belief, together with your assessment of your peer's belief as insane. Again, I don't see that this bit of first-personal information breaks the symmetry. For we can easily restore symmetry by adding that your peer returns the compliment by being extremely confident of her belief while finding yours insane (cf. Bogardus, pp. 329–30). Once you recognize this, there is no symmetry breaker anymore and you again lack reason to prefer your belief over hers.

Finally, Bogardus (2009, pp. 330–2) offers yet another variant of a first-personal symmetry breaker. He argues that *introspection* and *rational intuition* sometimes give us direct access to facts that we can "just see to be true". Sometimes you have "knowledge from direct acquaintance" that your own view is correct or that of your peer false. If you have such infallible knowledge, it will be rational to stick to your belief. This is why you shouldn't give in to your peer in the restaurant cases discussed by Christensen and Elga. Does this provide a symmetry breaker for internalists seeking to defend a steadfast view? Again, I think not. First, one might be skeptical about the possibility of infallible knowledge as such. Fallibilism for all kinds of knowledge is widely accepted in contemporary epistemology (Dougherty 2010). With the possible exceptions of elementary mathematical, logical, and conceptual truths and truths about the contents of our own consciousness — topics about which genuine peer disagreement seems impossible or unlikely anyway — it seems implausible that humans could possess infallible knowledge about anything. Given fallibilism, disagreement with a

known peer is exactly the kind of thing that makes the possibility of error salient and should thus lead you to question your initial belief. At best, then, this suggestion will work for a very small subclass of disagreements, to wit those that concern a topic about which we can obtain infallible knowledge and could conceivably disagree with a peer. Second, however, even if there are truths we can know infallibly, it's not clear that Bogardus's proposal can do much for the rationality *internalist*, who can only help herself to factors that are available from within her cognitive perspective.[17] Consider what an act of infallibly knowing something by introspection or rational intuition could look like from within your cognitive perspective. Presumably, you would form a belief and feel certain of its truth; it would seem to you that you just see that the belief is true and that you couldn't possibly be mistaken. But, crucially, *the fact that you infallibly know something* remains outside your cognitive perspective. Even if they exist, instances of infallible knowledge will presumably not come with a special mental halo to single them out from lesser kinds of knowledge.[18] Internalists, then, can only appeal to infallible knowledge in so far as it is accompanied by internal markers like extreme confidence, feelings of certainty, and similar cognitively accessible phenomena. These markers, however, could again be entirely symmetrical in the disagreeing parties. Alice could infallibly know that p and Beth mistakenly believe that not-p, yet both could experience the same confidence, certainty, etc. and sincerely report this to each other, so that their situations would again end up being entirely symmetrical.[19] The same considerations as before then lead to a council of conciliation.[20]

In sum, then, first-personal perspectives on the evidence do not break the symmetry in any relevant way. Since this exhausts the options for internalists, the conclusion from the previous section is vindicated: Internalists cannot be steadfast.

6. Externalism and Steadfastness

The argument so far is incomplete. Perhaps the reason that internalists cannot defend a steadfast view in cases of idealized disagreement is that *nobody* can. So the third step of my argument is to show that this is not the case. Externalism does have the resources to defend the rationality of a steadfast response.[21] One clarification: I aim to show that some forms of externalism *that are reasonably plausible on independent grounds* can support a steadfast response to idealized disagreement, not merely that it is possible to gerrymander an externalist account of rationality, which supports a steadfast response. Doing the latter would be too easy. For instance, if rational belief is equated with reliably formed *true* belief, only one of the disagreeing parties will have a rational belief in the first place, to which she should then probably hold on. An account that has truth immediately built into the conditions for rationality, however, is hardly defensible, since everyone agrees that there are instances of rationally held but false beliefs.

I'll briefly introduce one paradigmatic form of externalism to explain how the problem of disagreement is most naturally construed for externalists and then show how resources from various strands of the literature on externalism can be employed to defend a steadfast position. Alvin Goldman (1979) initially proposed a simple and robust form of externalism, *process reliabilism*. According to it, a belief is rational iff it has been produced by a reliable cognitive process, i.e., a process that produces a high ratio of true over false beliefs (in normal circumstances in the actual world).

It might seem that a steadfast response follows directly from this construal of rationality. In the idealized case of disagreement we're considering both Alice and Beth use a reliable process to form their respective beliefs and both therefore have a rational belief. Learning that their peer disagrees doesn't change this fact. Therefore both Alice and Beth should stick to their beliefs. But this line of reasoning overlooks a crucial consideration:

Forming beliefs through testimony from a reliable source is also a reliable cognitive process. So when Alice learns about Beth's contrary belief, she ends up with two conflicting beliefs, both of which have been formed by a reliable process. This conflict must be handled.[22]

The default move is to augment the basic account of rationality with a no-defeater clause to the effect that only beliefs that do not have undefeated defeaters are rational (Goldman 1986, p. 112f.; Bergmann 2006, pp. 153–77; Greco 2010, pp. 156–73). Following Bergmann (2005, pp. 422–4), we can understand a defeater as an experience or propositional attitude that one comes to have and that takes away the rationality of one's initial belief. Defeaters come in two kinds: a *rebutting* defeater is a ground or reason to think that your initial belief is false and an *undercutting* defeater is a ground or reason to think that the grounds or reasons for your initial belief are not indicative of its truth.[23] Peer disagreement presents a defeater, at least at first sight. Since your peer's belief contradicts yours, the defeater is most naturally construed as a rebutting defeater. But by making salient the possibility that you made a mistake in forming your initial belief, disagreement perhaps also casts doubt on the reliability of your belief-forming process, thus constituting an undercutting defeater.

So disagreement does pose a challenge to externalist accounts of rationality by presenting a defeater. The crucial question is whether this defeater can (sometimes) be defeated. If so, a steadfast response becomes possible. It is precisely here that externalists can appeal to external factors that are unavailable to internalists to break the symmetry. They can urge that, even though there are no internally accessible asymmetries between the two peers, it is nonetheless the case that asymmetrical factors *external* to the peers' cognitive perspectives make one of them rational in holding on to her belief and the other not. To see how this could go, I'll sketch some options. My goal is not to develop a detailed defense of any of these options; I merely want to show how externalism contains resources for doing so.

Externalists who explain rationality in terms of *proper cognitive functioning* (Bergmann 2006) can argue that, in some cases of peer disagreement, proper functioning requires that you hold on to your belief. Externalists who explain rationality in terms of *virtuous belief-formation* (Sosa 2007; Greco 2010) can argue that a virtuous cognitive agent would in fact be disposed to hold on to her own belief in some cases of disagreement. In such cases, disagreement doesn't give you an undefeated defeater after all. Of course, this raises questions: Why would proper functioning sometimes require steadfastness and why would a virtuous agent sometimes be disposed steadfastly? The externalist could answer that no further systematic arguments can be given to demonstrate that this is indeed what proper functioning or virtuous belief-formation requires. The most she can do to convince others is to present suitable cases in which one is supposed to *see* that this is how things are. This, for instance, is the line Bergmann (2009, p. 345) takes.[24]

Even if this position is correct, there is something unsatisfactory about not having some argument for why a steadfast response is correct in some cases. Alternatively, then, the externalist can make an effort to work out the mechanics of cases in which disagreement does not give you a defeater for your initial belief. Kelly (2010) serves as a good example of this. On the view he favors, the Total Evidence View, the peer who has in fact evaluated the first-order evidence (i.e., the evidence apart from the disagreement) correctly will typically be in a stronger epistemic position than her peer, who mistakenly thinks she has evaluated the first-order evidence correctly. Specifically, she will have a better justification than her peer for thinking that the first-order evidence supports her view. This is because correctly responding to the evidence creates 'upward epistemic push' (ibid., p. 159): It gives you justification for also believing that your evidence supports your belief. For the peer in the stronger epistemic position, a steadfast response is rational.

Now note that the factor creating the asymmetry between the peers is *who has in fact evaluated the first-order evidence correctly*. Crucially, this is not something that is accessible form within the peers' cognitive perspectives, as Kelly himself admits when he writes that "as a phenomenological matter, there might be no introspectible difference between how things seem when one is responding correctly and how things seem when one is not" (ibid., p. 169). This is really to say an *external* factor determines the rationality of the peers' responses to disagreement, which makes Kelly's view an externalist one according to the definition given in section 3.[25]

Lackey's (2010) 'justificationist view' provides another example. Although she doesn't intend her view to apply to the idealized cases of disagreement I have been considering, it can easily be co-opted to this end. The core of her view is that a steadfast response is in order in cases where one peer's initial belief is highly justified *and* she has a symmetry breaker in the form of 'personal information' about the cognitive processes she used to arrive at your original belief. This personal information can comprise evidence from introspection or intuition about the genuine quality of your evidence, the strength of your justification, or the correctness of your belief. I stand by my earlier conclusion that such personal information, when understood along internalist lines, cannot provide a good symmetry breaker in cases of idealized disagreements, but the important point here is that Lackey explicitly says that she construes justification externalistically, as requiring the reliability of the belief-forming process in question (ibid., p. 320). Since having a highly justified original belief is one of the factors that accounts for the rationality of a steadfast response, we again find that an external factor provides a relevant symmetry breaker.[26]

In conclusion, then, externalist accounts of rationality indeed make it possible to defend a steadfast response, even in idealized cases of peer disagreement. And that completes the third and final step of my argument.

7. Conclusion

Epistemic rationality internalism leads to conciliationism in idealized cases of peer disagreement. Since both peers must grant that their situations are entirely symmetrical and that it is impossible to obtain a symmetry breaker from within their own cognitive perspectives, each should adjust her belief in the direction of the other. Not doing so would be arbitrary. Rationality externalism does make it possible to defend a steadfast response to idealized peer disagreement. In essence, this is because externalists, unlike internalists, can appeal to factors having to do with the *de facto* asymmetry between the peers' opposing beliefs, which are unavailable from within their cognitive perspectives. Only externalists can thus be steadfast.

Acknowledgements

I'm indebted to the audience at the European Epistemology Network meeting in Lund in March 2010 and to the members of the Netherlands Research Seminar for Analytic Philosophy for helpful discussions about ancestors of this paper. Special thanks to Brent Madison for pointing me to relevant literature, to René van Woudenberg for feedback on an earlier version of this paper, to an anonymous referee of this journal for very generous and constructive comments, and to the editors of this special issue for their patience. Research for this article was supported by a Veni grant (275-20-023) from the Netherlands Organization for Scientific Research (NWO).

References

Alston, William P. (1988). An Internalist Externalism. *Synthese*, 74(3), 265–283.

Audi, Robert (2003). Epistemology: A Contemporary Introduction to the Theory of Knowledge. 2nd ed. (London: Routledge)

Ballantyne, Nathan (2011). Acquaintance and Assurance. Philosophical Studies (in press). DOI 10.1007/s11098-011-9747-9

——— & Coffman, E.J. (2011a). Uniqueness, Evidence, and Rationality. Philosophers' Imprint, 11(18), 1–13. http://quod.lib.umich.edu/cgi/p/pod/dod-idx?c=phimp;idno=3521354.0011.018.

———. (2011b). Conciliationism and Uniqueness. Australasian Journal of Philosophy (in press). DOI 10.1080/00048402.2011.627926

Bergmann, Michael (2005). Defeaters and Higher-Level Requirements. The Philosophical Quarterly, 55(220), 419–436.

——— (2006). Justification Without Awareness: A Defense of Epistemic Externalism. (New York: Oxford University Press)

——— (2009). Rational Disagreement After Full Disclosure. Episteme, 6(3), 336–353.

Bernecker, Sven & Pritchard, Duncan (Eds.) (2010). The Routledge Companion to Epistemology. (London: Routledge)

Bogardus, Tomas (2009). A Vindication of the Equal-Weight View. Episteme, 6(3), 324–335.

BonJour, Laurence (2010a). Externalism/Internalism. (In Jonathan Dancy, Ernest Sosa, & Matthias Steup (Eds.), A Companion to Epistemology (pp. 364–368). Oxford: Wiley-Blackwell.)

——— (2010b). Recent Work on the Internalism-Externalism Controversy. (In Jonathan Dancy, Ernest Sosa, & Matthias Steup (Eds.), A Companion to Epistemology (pp. 33–43). Oxford: Wiley-Blackwell.)

Christensen, David (2007). Epistemology of Disagreement: The Good News. Philosophical Review, 116(2), 187–217.

——— 2009. Disagreement as Evidence: The Epistemology of Controversy. Philosophy Compass, 4(1), 1–12.

——— 2011. Disagreement, Question-Begging and Epistemic Self-Criticism. Philosophers' Imprint, 11(6), 1–22.

Dougherty, Trent (2010). Fallibilism. (In Bernecker & Pritchard, pp. 131–143.)

Elga, Adam (2007). Reflection and Disagreement. Noûs, 41(3), 478–502.

Enoch, David (2010). Not Just a Truthometer: Taking Oneself Seriously (but Not Too Seriously) in Cases of Peer Disagreement. Mind (in press). DOI 10.1093/mind/fzq070

Feldman, Richard (2006). Epistemological Puzzles About Disagreement. (In Stephen Hetherington (Ed.), Epistemology Futures (pp. 216–236). Oxford: Oxford University Press.)

——— (2007). Reasonable Religious Disagreements. (In Louise Antony (Ed.), Philosophers Without Gods (pp. 194–214). New York: Oxford University Press.)

——— & Earl Conee (2001). Internalism Defended. American Philosophical Quarterly, 38(1), 1–18.

——— & Ted Warfield (Eds.) (2010). Disagreement. (New York: Oxford University Press)

Frances, Bryan (2010a). The Reflective Epistemic Renegade. Philosophy and Phenomenological Research, 81(2), 419–463.

——— (2010b). Disagreement. (In Bernecker & Pritchard, pp. 68–74.)

Fumerton, Richard (1995). Metaepistemology and Skepticism. (Lanham, MD: Rowman & Littlefield)

——— (2006). Epistemology. (Oxford: Blackwell)

——— (2010). Poston on Similarity and Acquaintance. Philosophical Studies, 147(3), 379–386.

Goldman, Alvin (1979). What Is Justified Belief? (In George Pappas (Ed.), Justification and Knowledge (pp. 1–23). Dordrecht: Reidel.)

——— (1986). Epistemology and Cognition. (Cambridge, MA: Harvard University Press)

Greco, John (2010). Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity. (Cambridge: Cambridge University Press)

Haddock, Adrian, Alan Millar, & Duncan Pritchard (Eds.) (2010). Social Epistemology. (Oxford: Oxford University Press)

Kelly, Tom (2005). The Epistemic Significance of Disagreement. (In John Hawthorne & Tamar Szabo Gendler (Eds.), Oxford Studies in Epistemology, Vol. I (pp. 167–196). Oxford: Oxford University Press).

——— (2010). Peer Disagreement and Higher-Order Evidence. (In Feldman & Warfield, pp. 111–174.)

King, Nathan L. (2011). Disagreement: What's the Problem? Or: a Good Peer Is Hard to Find. Philosophy and Phenomenological Research (in press). DOI 10.1111/j.1933-1592.2010.00441.x

Lackey, Jennifer (2010). A Justificationist View of Disagreement's Epistemic Significance. (In Haddock et al., pp. 298–325.)

Meeker, Kevin (2004). Justification and the Social Nature of Knowledge. Philosophy and Phenomenological Research, 69(1), 156–172.

Pollock, John L. (1974). Knowledge and Justification. (Princeton, NJ: Princeton University Press)

Poston, Ted (2010). Similarity and Acquaintance: A Dilemma. Philosophical Studies, 147(3), 369–378.

Pritchard, Duncan (2010). What Is This Thing Called Knowledge? 2nd ed. (London: Routledge)

Sosa, Ernest (2007). A Virtue Epistemology: Apt Belief and Reflective Knowledge, Vol. I. (New York: Oxford University Press)

——— (2010). The Epistemology of Disagreement. (In Haddock et al., pp. 278–297.)

Steup, Matthias (1996). An Introduction to Contemporary Epistemology. (Upper Saddle River, NJ: Prentice Hall)

Van Inwagen, Peter (1996). It Is Wrong Everywhere, Always, and for Anyone, to Believe Anything Upon Insufficient Evidence. (In Jeff Jordan and Daniel Howard-Snyder (Eds.), Faith, Freedom, and Rationality (pp. 137–153). Lanham, MD: Rowman & Littlefield.)

——— (2010). We're Right. They're Wrong. (In Feldman & Warfield, pp. 10–28.)

Wedgwood, Ralph (2010). The Moral Evil Demons. (In Feldman & Warfield, pp. 216–246.)

White, Roger (2005). Epistemic Permissiveness. Philosophical Perspectives, 19, 445–459.

——— (2009). On Treating Oneself and Others as Thermometers. Episteme, 6(3), 233–250.

[1] Ballantyne & Coffman (2011a) are an exception. They show how views on evidence and rationality distilled from earlier epistemological literature bear on the Uniqueness Thesis (White 2005; Feldman 2007, p. 205).

[2] Feldman (2006) adds a third condition to the effect that peers must have disclosed their evidence fully. This is fine, but (i) can also be read as already entailing full disclosure.

[3] Fumerton (2006) does the same. See Pritchard (2010, pp. 49–51) for a brief discussion of the issues involved in identifying rational with justified belief and Audi (2003) for an account that construes them differently.

[4] For similar characterizations, see Steup (1996, p. 84), Audi (2003, pp. 240–2). Feldman and Conee (2001) offer an alternative account of internalism as *mentalism*: the thesis that all the factors determining rationality are *mental*, i.e., 'internal to the person's mental life', but not necessarily accessible upon reflection. For arguments that this account of internalism is misguided, see Bonjour (2010b, pp. 34–6) and Bergmann (2006, pp. 45–75).

[5] Such conditions are typically not stated explicitly, but they strike me as sensible additions. It's not so clear what to make of situations in which both peers are dimwitted and epistemically irresponsible, or when they hold beliefs about issues wide outside their areas of competence.

[6] The defenses of conciliationism in Christensen (2007), Elga (2007), and Bogardus (2009) also hinge on considerations of symmetry.

[7] I borrow this term from Lackey (2010, p. 309), who credits Nathan Christiansen for coming up with it.

[8] On this thermometer analogy, see White (2009). Enoch (2010) even claims the thermometer analogy underlies all defenses of conciliationism but then argues forcefully that the analogy is misguided.

[9] Elga (2007, pp. 488–91) stresses the importance for conciliationism of taking into account such "circumstances of the disagreement."

[10] See Frances (2010b, pp. 71–4) for a brief exploration of the epistemic significance of consensus among a group of peers or superiors.

[11] For this reason, Kelly (2010, pp. 138–9) retracts his earlier position.

[12] Christensen (2007; 2011) and Elga (2007) explicitly endorse such an independence constraint.

[13] Christensen (2011), however, argues that this is not so. He claims that Kelly's and Lackey's cases fail to constitute counterexamples to a properly understood independence constraint.

[14] It is hard to give uncontroversial examples of this, but one could think of the impression an argument makes on you or intuitions you have about philosophical positions. E.g., an argument or position striking you as feeble, even while you're unable (so far) to pinpoint any objective problems with it.

[15] To claim that having different first-personal evaluations of the evidence is incompatible with being peers makes the conditions of peerhood virtually impossible to satisfy. Such a move would thus define the problem of *peer* disagreement away, although most of the same questions would then resurface as questions about the rational response to disagreement with near-peers.

[16] For those who doubt that this makes sense, I would argue that if the case was possible in the first place, then this should be possible as well.

[17] What follows isn't necessarily a criticism of Bogardus. Since he doesn't say whether his proposal should be understood along internalist or externalist lines, it could be that he is really thinking along externalist lines, in which case the proposal may well work, as I'll suggest below (cf. note 26).

[18] Unless you endorse a version of the KK principle from which it follows that infallible knowledge automatically gives you (infallible?) second-order knowledge that you know infallibly, which seems implausible *in excelsis*.

[19] I will readily admit that this is implausible — perhaps impossible — for cases where the disagreement concerns an area where infallible knowledge may be possible. Viz., could Alice really feel certain that 5 + 7 = 12 because she infallibly knows it, while Beth, who is supposed to be her peer, feels equally certain that 5 + 7 = 13? Then again, as I said above about Christensen's restaurant case, it is implausible that there could be genuine *peer* disagreement in such cases in the first place (cf. also note 16 above). Once we've accepted the latter possibility, however, the former should also be granted.

[20] An anonymous referee pointed out that Fumerton's acquaintance theory of noninferential justification (Fumerton 1995) could be employed to buttress Bogardus's proposal, at least when Fumerton's theory is interpreted along infallibilist lines (which is what Fumerton has sometimes had in mind (cf. Poston 2010), even though he has recently argued in favor of a fallibilist interpretation (Fumerton 2010)). On this account, for a subject to have noninferential justification for a belief that p, the subject must be directly acquainted with her belief that p, the fact that p, and the correspondence relation between the fact and the belief. Since acquaintance entails awareness, and hence cognitive access, the theory is clearly internalist. Now suppose Alice and Beth disagree about whether p. As it turns out, Alice has infallible noninferential justification for her belief that p, while Beth merely thinks that she does. Since only Alice, and not Beth, is in fact directly acquainted with the fact that p and the correspondence between that fact and her belief, it will not be the case that all their internal markers are exactly the same. Alice has a relevant symmetry breaker. Hence, we have an internalist account of rationality (justification) that allows one to defend a steadfast response to disagreement with a known peer.

While a full discussion of Fumerton's views is beyond the pale here, I will say three things in response. First, if there indeed is infallible noninferential justification, it will be for propositions about which genuine peer disagreement seems impossible. Perhaps I can be infallibly justified in believing that I am in intense pain, but that is not something about which I can disagree with a peer, since the evidence is crucially first-personal. Perhaps I can be infallibly justified that 5 + 7 = 12 (but: what would it mean to be 'directly acquainted' with the fact that 5 + 7 = 12?), but again it seems that there couldn't be a disagreeing peer, since any educated and healthy adult who understands the relevant concepts will agree that 5 + 7 = 12. For any case about which genuine peer disagreement is conceivable, it seems our justification will be fallible (even if it is noninferential). Second, I'm not sure that sense can be made of the suggestion that the *internal* markers will really be

asymmetrical in a case where Alice has infallible noninferential justification and Beth sincerely and confidently thinks she does but in fact doesn't. As I argued above in response to Bogardus's proposal, considered purely from within one's cognitive perspective, an instance of infallibly knowing that p will presumably look exactly like an instance of *thinking* that one infallibly knows that p. It's easy to see how markers like certainty, confidence, and believing that you couldn't possibly be mistaken are internally accessible and distinguishable from their counterparts, but what could it be for actual acquaintance with the fact that p to be both accessible *and distinguishable* from merely apparent direct acquaintance with an alleged fact that not-p? Third, even if Alice has infallible noninferential justification and is thus directly aware of the fact that p and the correspondence of that fact with her belief that p, it seems that her peer Beth's reporting that she believes that not-p should at least lead Alice to consider the possibility that she is mistaken. To assure herself that she isn't, she would have to acquire a higher-order justified belief that she is infallibly justified in her first-order belief that p. Apart from the fact that we see the beginnings of a vicious regress here (cf. Ballantyne 2011 for a similar regress argument against Fumerton's view), this leads to a dilemma. (a) If there is only fallible (inferential or noninferential) justification for this second-order belief, the possibility that she is mistaken will remain salient for Alice and it would be arbitrary for her to hold on to her belief. (b) To claim that there is also infallible (noninferential) justification for this second-order belief, however, seems exceedingly implausible. Alice would have to be directly acquainted with the fact that she has infallible justification for her belief that p and of the correspondence relation between this fact and her belief that she is so justified in believing that p. It doesn't seem that second-order facts about our justification and their relations with our second-order beliefs are the sorts of things we are directly acquainted with, nor is it clear that such facts could be accessible from within our cognitive perspectives. At any rate, we can conclude that internalists who want to defend a steadfast response to disagreement are driven towards a extremely demanding and wildly implausible conception of infallible justification. I think it's safe to take that as support for the overall argument of this paper.

[21] Ballantyne & Coffman (2011a) argue that rationality externalism, when combined with most prominent conceptions of the nature of evidence, entails the falsity of the Uniqueness Thesis. Conciliationism is often taken to carry with it a commitment to Uniqueness (Kelly 2010, p. 121; Ballantyne & Coffman 2011b). This suggests that if Uniqueness is false, it should be easier to defend a non-conciliationist (steadfast) view. Since I intend to show that rationality externalism (which — if Ballantyne & Coffman are right — should lead many of its proponents to reject Uniqueness) indeed allows one to defend a steadfast response, our conclusions are complementary. I only learned about their paper after I had written most of the present paper.

[22] The same problem arises on other crude forms of externalism, on which rational belief is construed as being the result of properly functioning cognitive mechanisms or of virtuous belief-formation.

[23] This distinction goes back at least to Pollock (1974, pp. 41–3), where they are called type I and type II defeaters respectively.

[24] Greco (2010, pp. 167–73) also claims that in some cases a virtuous agent will in fact respond to a defeater by giving up her belief, whereas in others she won't, without attempting to give an account of the differences between the two kinds of cases that warrant these responses.

[25] It is remarkable that Kelly himself explicitly disavows the externalist label (2010, p. 169). He argues that his view resembles internalist views in an important respect: "On the Total Evidence View, what it is reasonable for one to believe always depends on one's total evidence, and only considerations of which one is aware are eligible for inclusion in one's total evidence" (ibid.). I cannot help but think that this is confused. This gloss on his own view omits the crucial fact that what it is ultimately reasonable for the peers to believe *also* depends on who has *in fact* responded correctly to the evidence. And that, as he himself grants, is not a consideration of which one is aware.

[26] Yet another option would be to interpret Bogardus's (2009) view, discussed above, in an explicitly externalist fashion as claiming that the peer who has *in fact* introspected or intuited correctly that she is right or her peer wrong will be rational to hold on to her belief. As I said above, however, I think this view can only offer solace for a very limited subclass of disagreements concerning topics about which peer disagreement is unlikely in the first place.